

Early Journal Content on JSTOR, Free to Anyone in the World

This article is one of nearly 500,000 scholarly works digitized and made freely available to everyone in the world by JSTOR.

Known as the Early Journal Content, this set of works include research articles, news, letters, and other writings published in more than 200 of the oldest leading academic journals. The works date from the mid-seventeenth to the early twentieth centuries.

We encourage people to read and share the Early Journal Content openly and to tell others that this resource exists. People may post this content online or redistribute in any way for non-commercial purposes.

Read more about Early Journal Content at http://about.jstor.org/participate-jstor/individuals/early-journal-content.

JSTOR is a digital library of academic journals, books, and primary source objects. JSTOR helps people discover, use, and build upon a wide range of content through a powerful research and teaching platform, and preserves this content for future generations. JSTOR is part of ITHAKA, a not-for-profit organization that also includes Ithaka S+R and Portico. For more information about JSTOR, please contact support@jstor.org.

THE ANALYST.

VOL. III. NOVEMBER, 1876. No. 6.

NEW INVESTIGATION OF THE LAW OF ERRORS OF OBSERVATION.

BY CHAS. H. KUMMELL, ASSISTANT U. S. LAKE SURVEY, DETROIT, MICH.

(*Continued from page* 140.)

5. If m observations are made of a quantity M which is a known function of n unknown quantities $x, y, z, \ldots w$ we have a determinate problem if $m = n$. However if $m > n$ as we have supposed, the problem admits of an infinite number of contradictory systems, since each observation is liable to error. It is a logical necessity to remove these contradictions and this is done by applying corrections to the observed quantities $M_1, M_2 \ldots . M_m$, which may be $\Delta_1, \Delta_2, \ldots . \Delta_m$, respectively. Since we have no rigorous condition to determine these corrections we may ask for the most probable system of corrections and this problem may pe solved by the principles already explained.

Assuming the observations to be of equal precision, (if not, they may be reduced to the same precision by (26) if the precisions are given or some numbers proportional to them,) we have, if the probable error of observation is r, the compound probability of the errors $\Delta_1, \Delta_2, \ldots . \Delta_m$, to be committed in any order

$$P = \varphi_0^m e^{-\frac{\rho^2}{r^2}[\triangle^2]}. \tag{44}$$

If this probability shall be a maximum then

$$[\Delta^2]_0 = \text{a minimum}, \tag{45}$$

if the subscript 0 denotes that the corrections already satisfy the condition.

Now let M be a linear function of $x, y, z, \ldots . w$,[if not we can prepare it as (34″),] or let M be of the form

$$M = ax + by + cz + \ldots . lw + k, \tag{46}$$

then we have for each observation, replacing $M_1 - k$, $M_2 - k$, ... $M_m - k$ by $\mu_1, \mu_2, \ldots \mu_m$, and supplying corrections

$$
\begin{aligned}
(47) \qquad & a_1x + b_1y + c_1z + \ldots + l_1w - \mu_1 = \Delta_1 \\
& a_2x + b_2y + c_2z + \ldots + l_2w - \mu_2 = \Delta_2 \\
& \ldots\ldots\ldots\ldots\ldots\ldots \\
& a_mx + b_my + c_mz + \ldots + l_mw - \mu_m = \Delta_m
\end{aligned}
$$

By (45) we must have

$$
\frac{d[\Delta^2]_0}{dx} = \frac{d[\Delta^2]_0}{dy} = \frac{d[\Delta^2]_0}{dz} = \ldots \frac{d[\Delta^2]_0}{dw} = 0, \qquad \text{or}
$$

$$
\begin{aligned}
(48) \qquad & \Delta_1\frac{d\Delta_1}{dx} + \Delta_2\frac{d\Delta_2}{dx} + \ldots \Delta_m\frac{d\Delta_m}{dx} = 0 \\
& \Delta_1\frac{d\Delta_1}{dy} + \Delta_2\frac{d\Delta_2}{dy} + \ldots \Delta_m\frac{d\Delta_m}{dy} = 0 \\
& \ldots\ldots\ldots\ldots\ldots \\
& \Delta_1\frac{d\Delta_1}{dw} + \Delta_2\frac{d\Delta_2}{dw} + \ldots \Delta_m\frac{d\Delta_m}{dw} = 0,
\end{aligned}
$$

but by (47)

$$
\begin{aligned}
(49) \qquad & \frac{d\Delta_1}{dx} = a_1; \ \frac{d\Delta_2}{dx} = a_2 \ldots \frac{d\Delta_m}{dx} = a_m \\
& \frac{d\Delta_1}{dy} = b_1; \ \frac{d\Delta_2}{dy} = b_2 \ldots \frac{d\Delta_m}{dy} = b_m \\
& \ldots\ldots\ldots\ldots\ldots \\
& \frac{d\Delta_1}{dw} = l_1; \ \frac{d\Delta_2}{dw} = l_2 \ldots \frac{d\Delta_m}{dw} = l_m
\end{aligned}
$$

These values substituted in (48) give

$$
\begin{aligned}
(50) \qquad & a_1\Delta_1 + a_2\Delta_2 + \ldots a_m\Delta_m = [a\Delta] = 0 \\
& b_1\Delta_1 + b_2\Delta_2 + \ldots b_m\Delta_m = [b\Delta] = 0 \\
& \ldots\ldots\ldots\ldots\ldots \\
& l_1\Delta_1 + l_2\Delta_2 + \ldots l_m\Delta_m = [l\,\Delta] = 0.
\end{aligned}
$$

Multiplying (47) in their order by $a_1, a_2 \ldots a_m$ and adding the products, then also taking the sum of the products by $b_1, b_2, \ldots b_m$ and so on we obtain regarding (50)

$$
\begin{aligned}
(51) \quad & [a^2]x + [ab]y + [ac]z + \ldots + [al]w - [a\mu] = [a\Delta] = 0 \\
& [ab]x + [b^2]y + [bc]z + \ldots + [bl]w - [b\mu] = [b\Delta] = 0 \\
& [ac]x + [bc]y + [c^2]z + \ldots + [cl]w - [c\mu] = [c\Delta] = 0 \\
& \ldots\ldots\ldots\ldots\ldots\ldots \\
& [al]x + [bl]y + [cl]z + \ldots + [l^2]w - [l\,\mu] = [l\,\Delta] = 0.
\end{aligned}
$$

These equations which are called normal equations will give those determinate values of x, y, z which are rendered most probable by the observations.

Multiplying (47) in succession by $-\mu_1, -\mu_2, \ldots -\mu_m$ and adding the products we have also

$$-[a\mu]x - [b\mu]y - [c\mu]z - \ldots - [l\mu]w + [\mu^2] = -[\mu\Delta].$$

and since the sum of the products by $\Delta_1, \Delta_2, \ldots \Delta_m$

$$[a\Delta]x + [b\Delta]y + [c\Delta]z + \ldots + [l\Delta]w - [\mu\Delta] = -[\mu\Delta] = [\Delta^2]_0$$

we have

(52) $$-[a\mu]x - [b\mu]y - [c\mu]z - \ldots - [l\mu]w + [\mu^2] = [\Delta^2]_0.$$

Having found $x, y, z, \ldots w$ by (51) we can find the sum of minimum squares of the corrections $[\Delta^2]_0$ by forming also the quantity $[\mu^2]$. We have then by (23′)

$$\epsilon_\Delta = \sqrt{\left(\frac{[\Delta^2]_0}{m}\right)},$$

and by (30) the probable error of the corrections

(53) $$r_\Delta = \rho\sqrt{\left(\frac{2[\Delta^2]_0}{m}\right)}.$$

This formula is of course only approximately true and rests on the assumption that the system of corrections $\Delta_1, \Delta_2, \ldots \Delta_m$ sufficiently near corresponds to the ideal case in article 2. This must be the case as nearly as possible since it is the most probable system of corrections.

6. We next determine the probable errors of $x, y, z, \ldots w$. It is plain that these unknown quantities must be some linear functions of $\mu_1, \mu_2, \ldots \mu_m$, so that we may assume

(54)
$$\begin{aligned} x &= \alpha_1\mu_1 + \alpha_2\mu_2 + \ldots \alpha_m\mu_m = [\alpha\mu] \\ y &= [\beta\mu] \\ z &= [\gamma\mu] \\ &\ldots \\ w &= [\lambda\mu]. \end{aligned}$$

The probable error of observation or that of $\mu_1, \mu_2, \ldots \mu_m$ being r, we have by (43′)

(55)
$$\begin{aligned} r_x &= r\sqrt{[\alpha^2]} \\ r_y &= r\sqrt{[\beta^2]} \\ r_z &= r\sqrt{[\gamma^2]} \\ &\ldots \\ r_w &= r\sqrt{[\lambda^2]}. \end{aligned}$$

Substituting (54) into (47), (50) and (52) we have

$$(47') \quad \begin{aligned} &a_1[a\mu] + b_1[\beta\mu] + c_1[\gamma\mu] + \ldots + l_1[\lambda\mu] - \mu_1 = \Delta_1 \\ &a_2[a\mu] + b_2[\beta\mu] + c_2[\gamma\mu] + \ldots + l_2[\lambda\mu] - \mu_2 = \Delta_2 \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ &a_m[a\mu] + b_m[\beta\mu] + c_m[\gamma\mu] + \ldots + l_m[\lambda\mu] - \mu_m = \Delta_m \end{aligned}$$

$$(50') \quad \begin{aligned} &[a^2][a\mu] + [ab][\beta\mu] + [ac][\gamma\mu] + \ldots [al][\lambda\mu] - [a\mu] = 0 \\ &[ab][a\mu] + [b^2][\beta\mu] + [bc][\gamma\mu] + \ldots [bl][\lambda\mu] - [b\mu] = 0 \\ &[ac][a\mu] + [bc][\beta\mu] + [c^2][\gamma\mu] + \ldots [cl][\lambda\mu] - [c\mu] = 0 \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ &[al][a\mu] + [bl][\beta\mu] + [cl][\gamma\mu] + \ldots [l^2][\lambda\mu] - [l\mu] = 0. \end{aligned}$$

$$(52') \quad -[a\mu][a\mu] - [b\mu][\beta\mu] - [c\mu][\gamma\mu] - \ldots - [l\mu][\lambda\mu] - [\mu^2] = [\Delta^2]_0.$$

The coefficients $a_1, a_2 \ldots a_m$; $\beta_1, \beta_2, \ldots$ being independent of $\mu_1, \mu_2, \ldots \mu_m$ we have, resolving (50′),

$$(56) \quad \begin{aligned} &[a^2]a_1 + [ab]\beta_1 + [ac]\gamma_1 + \ldots [al]\lambda_1 = a_1 \\ &[a^2]a_2 + [ab]\beta_2 + [ac]\gamma_2 + \ldots [al]\lambda_2 = a_2 \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ &[ab]a_1 + [b^2]\beta_1 + [bc]\gamma_1 + \ldots [bl]\lambda_1 = b_1 \\ &[ab]a_2 + [b^2]\beta_2 + [bc]\gamma_2 + \ldots [bl]\lambda_2 = b_2 \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ &[al]a_1 + [bl]\beta_1 + [cl]\gamma_1 + \ldots [l^2]\lambda_1 = l_1 \\ &[al]a_2 + [bl]\beta_2 + [cl]\gamma_2 + \ldots [l^2]\lambda_2 = l_2. \\ &\cdots\cdots\cdots\cdots\cdots\cdots \end{aligned}$$

Taking sums of products by $a_1, a_2, \ldots$ then by $b_1, b_2, \ldots$ by $c_1, c_2, \ldots$ and so on, we easily see that we must have

$$(57) \quad \begin{aligned} &[a\alpha] = 1;\ [b\alpha] = 0;\ [c\alpha] = 0; \ldots [l\alpha] = 0 \\ &[a\beta] = 0;\ [b\beta] = 1;\ [c\beta] = 0; \ldots [l\beta] = 0 \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ &[a\lambda] = 0;\ [b\lambda] = 0;\ [c\lambda] = 0; \ldots [l\lambda] = 1. \end{aligned}$$

We have then, taking the sum of the products of (56) by $a_1, a_2, \ldots$

$$(58) \quad \begin{aligned} &[a^2][a^2] + [ab][a\beta] + [ac][a\gamma] + \ldots [al][a\lambda] = [aa] = 1 \\ &[ab][a^2] + [b^2][a\beta] + [bc][a\gamma] + \ldots [bl][a\lambda] = [ba] = 0 \\ &[ac][a^2] + [bc][a\beta] + [c^2][a\gamma] + \ldots [cl][a\lambda] = [ca] = 0 \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ &[al][a^2] + [bl][a\beta] + [cl][a\gamma] + \ldots [l^2][a\lambda] = [la] = 0. \end{aligned}$$

Multiplying the first equation of (51) by $[\alpha^2]$, the second by $[\alpha\beta]$, the third by $[\alpha\gamma]$ . . . . the nth by $[\alpha\lambda]$ and adding the products we have by virtue of (58)

$$(59)\quad x = [\alpha^2][a\mu] + [\alpha\beta][b\mu] + [\alpha\gamma][c\mu] + \ldots + [\alpha\lambda][l\mu].$$

In a similar manner we can determine y from the system

$$\begin{aligned}(58')\quad & [\alpha^2][\alpha\beta] + [ab][\beta^2] + [ac][\beta\gamma] + \ldots [al][\beta\lambda] = [a\beta] = 0\\ & [ab][\alpha\beta] + [b^2][\beta^2] + [bc][\beta\gamma] + \ldots [bl][\beta\lambda] = [b\beta] = 1\\ & [ac][\alpha\beta] + [bc][\beta^2] + [c^2][\beta\gamma] + \ldots [cl][\beta\lambda] = [c\beta] = 0\\ & \cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\\ & [al][\alpha\beta] + [bl][\beta^2] + [cl][\beta\gamma] + \ldots [l^2][\beta\lambda] = [l\beta] = 0.\end{aligned}$$

$$(59')\quad y = [\alpha\beta][a\mu] + [\beta^2][b\mu] + [\beta\gamma][c\mu] + \ldots + [\beta\lambda][l\mu],$$

and so on for the other unknown quantities.

To determine now the probable error of observation or that of $\mu_1, \mu_2, \ldots \mu_m$ we have from (47′), omitting subscripts,

$$\mu = a[a\mu] + b[\beta\mu] + c[\gamma\mu] + \ldots . l[\lambda\mu] - \Delta.$$

We have then the relation of probable error by (43′)

$$\left.\begin{aligned} r^2 = {} & (aa_1 + b\beta_1 + c\gamma_1 \ldots l\lambda_1)^2\\ & + (aa_2 + b\beta_2 + c\gamma_2 \ldots l\lambda_2)^2\\ & + \cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\\ & + (aa_m + b\beta_m + c\gamma_m \ldots l\lambda_m)^2\end{aligned}\right| r^2 + r_\Delta^2$$

$$\left.\begin{aligned} & = a^2[a^2] + ab[a\beta] + ac[a\gamma] + \ldots al[a\lambda]\\ & + ab[a\beta] + b^2[\beta_2] + bc[\beta\gamma] + \ldots bl[\beta\lambda]\\ & + \cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\\ & + al[a\lambda] + bl[\beta\lambda] + cl[\gamma\lambda] + \ldots l^2[\lambda^2]\end{aligned}\right| r^2 + r_\Delta^2$$

Restoring subscripts and taking sum of all these equations we have

$$\left.\begin{aligned} (60)\quad mr^2 = {} & [a^2][a^2] + [ab][a\beta] + [ac][a\gamma] + \ldots [al][a\lambda]\\ & + [ab][a\beta] + [b^2][\beta^2] + [bc][\beta\gamma] + \ldots [bl][\beta\lambda]\\ & + \cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\\ & + [al][a\lambda] + [bl][\beta\lambda] + [cl][\gamma\lambda] + \ldots [l^2][\lambda^2]\end{aligned}\right| r^2 + mr_\Delta^2$$

$$= nr^2 + mr_\Delta^2 \text{ by virtue of (58), (58′),}$$

We have then the remarkable relation

$$(61)\qquad r = r_\Delta\sqrt{\frac{m}{m-n}};$$

hence by (53)

$$(62)\qquad r = \rho\sqrt{\frac{2[\Delta^2]_0}{m-n}}.$$

This value being used in (55) we have also the probable errors of the unknown quantities.

7. If the observations are of different precision we must reduce them first to the same precision according to (26). Usually this is done however by means of weights, which are numbers proportional to the square of the precision, or what is the same thing proportional to the reciprocal square of the probable error. Thus if in last article we assign to an observation the weight unity those of $x, y, z, \ldots w$ are

$$(63) \qquad p_x = \frac{1}{[\alpha^2]}, \quad p_y = \frac{1}{[\beta^2]}, \quad \ldots \quad p_w = \frac{1}{[\lambda^2]}.$$

If then, as usual, weights are given we reduce the observations to the same precision if we multiply them by the square root of the weight and then proceed as in last article. Formula (62) gives then the probable error of that standard of observations to which they all have been reduced.

8. If there is only one unknown quantity x, (47) becomes

$$(64) \qquad \begin{aligned} a_1 x - \mu_1 &= \Delta_1 \\ a_2 x - \mu_2 &= \Delta_2 \\ &\cdots\cdots \\ a_m x - \mu_m &= \Delta_m \end{aligned}$$

These equations being of equal precision by hypothesis the coefficients a_1, a_2, a_m may be understood to be ratios of the actually observed quantity to the quantity sought, for instance, as if the quantity had been magnified in different degrees or repeated a different number of times. We have then from the first of (51)

$$(65) \qquad x = [a\mu] \div [a^2],$$

and from the first of (58)

$$(66) \qquad [\alpha^2] = 1 \div [a^2], \text{ or } p_x = [a^2].$$

We have also the probable error of observation by (62)

$$(67) \qquad r = \rho\sqrt{\frac{2[\Delta^2]_0}{m-1}};$$

hence

$$r_x = \rho\sqrt{\frac{2[\Delta^2]_0}{(m-1)[a^2]}}$$

If however we place

$$(69) \qquad \begin{aligned} x_1 &= \mu_1 \div a_1; \ \delta_1 = \Delta_1 \div a_1 \\ x_2 &= \mu_2 \div a_2; \ \delta_2 = \Delta_2 \div a_2 \\ &\cdots\cdots\cdots \\ x_m &= \mu_m \div a_m; \ \delta_m = \Delta_m \div a_m, \end{aligned}$$

we have the equations of condition

$$(70) \qquad \begin{aligned} x - x_1 &= \delta_1 && \text{(precision } a_1 \text{ or weight } a_1^2) \\ x - x_2 &= \delta_2 && (\text{“} \quad a_2 \text{“} \quad \text{“} \quad a_2^2) \\ &\cdots\cdots && \cdots\cdots \\ x - x_m &= \delta_m && (\text{“} \quad a_m \text{“} \quad \text{“} \quad a_m^2). \end{aligned}$$

We find after reducing to same precision

$$x = \frac{[a^2x]}{[a^2]} = \frac{[px]}{[p]}. \tag{71}$$

$$p_x = [p] \quad = \text{sum of weights of individual obs.} \tag{72}$$

$$r = \rho\sqrt{\frac{2[p\delta^2]}{m-1}} \tag{73}$$

$$r_x = \rho\sqrt{\frac{2[p\delta^2]}{(m-1)[p]}}. \tag{74}$$

Suppose now that $a_1 = a_2 = \ldots a_m$ or $p_1 = p_2 = \ldots p_m$, then the weighted mean (71) becomes the arithmetic mean and we have

$$x = [x] \div m \tag{71'}$$

$$p_x = m \tag{72'}$$

$$r = \rho\sqrt{\frac{2[\delta^2]}{m-1}} \tag{73'}$$

$$r_x = \rho\sqrt{\frac{2[\delta^2]}{(m-1)m}} \tag{74'}$$

9. The above short essay contains every essential principle of the method of least squares, though only a small part of the many interesting developments. I merely sketch out a method of treating the general problem of m observations for n unknown quantities. It consists in determining the probability of the system relatively to the unknown quantities.

We had (44)

$$P = \varphi^m e^{-\frac{\rho^2}{r^2}[\triangle^2]};$$

the probability of the system relatively to the unknown quantities is then

$$\Pi = \frac{P dx\, dy\, dz \ldots dw}{\int_{-\infty}^{+\infty} dw \ldots \int_{-\infty}^{+\infty} dz \int_{-\infty}^{+\infty} dy \int_{-\infty}^{+\infty} P\, dx} \tag{75}$$

and that of one unknown quantity for instance w

$$\varphi(w) = \frac{dw \int_{-\infty}^{+\infty} \ldots \int_{-8}^{+8} dz \int_{-\infty}^{+\infty} dy \int_{-\infty}^{+\infty} P\, dx}{\int_{-\infty}^{+\infty} dw \int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty} dz \int_{-\infty}^{+\infty} dy \int_{-\infty}^{+\infty} P\, dx}. \tag{76}$$